

REMARKS

This Amendment is submitted in response to the Office Action dated June 22, 2006, having a shortened statutory period set to expire September 22, 2006. The present amendment cancels Claim 13. No new subject matter is presented, and thus no new issues are raised by the present amendment. Claims 8, 10-12 and 15-20 are still pending.

Rejection under 35 U.S.C. §112

On page 2 of the present Office Action, the Examiner has rejected Claim 13 under 35 U.S.C. § 112, second paragraph, for the use of the term “or.” This claim is now cancelled, and thus the rejection is moot.

Rejections under 35 U.S.C. § 103

On page 3 of the present Office Action, Claims 8, 10-13 and 15-20 are rejected as being unpatentable over *Ishikawa et al.* (U.S. Patent No. 5,848,407 – “*Ishikawa*”) in view of *Gutierrez et al.* (U.S. Patent Application Publication No. 2003/0046276 – “*Gutierrez*”), and further in view of *Finseth et al.* (U.S. Patent No. 6,271,840 – “*Finseth*”). Applicants respectfully traverse these rejections.

Regarding *Gutierrez*, this art is not available under 35 U.S.C. 103(c) since the present application (evidenced by the assignment recorded on Reel 014642, Frame 0595 of the USPTO Assignment Division) and the cited art (on its face) are both assigned to International Business Machines Corporation of Armonk, NY. Thus, these rejections are invalid.

Even if *Gutierrez* were available as 103/102(e) art, the rejections should still fail for the following reasons.

With regards to exemplary Claim 8, a combination of the cited art does not teach or suggest that “unnecessary information elements are plural information elements that include an OBJECT_IMAGE having a same Uniform Resource Locator (URL), (as supported at paragraph

[0082]) wherein the OBJECT_IMAGE describes a type of media used to display the HTML document (as supported at paragraph [0059]).” The present Office Action cites *Finseth* for teaching this feature in the abstract and at col. 5, lines 32-35 and lines 62-67. *Finseth*, and in particular at the cited passage, does not teach or suggest this feature. *Finseth* never teaches or suggests multiple descriptors (of a same type of media) that have a same URL. Rather, the cited passages refer to rendering web pages that are called up by URLs. There is no teaching or suggestion of defining OBJECT_IMAGES that have a same URL as being “unnecessary information elements.”

With regards to independent Claim 19, a combination of the cited art does not teach or suggest a method for eliminating ambiguity of a specified topic being searched during a web crawling, which is supported in the present specification at paragraph [0189], and includes the steps of: “presenting relevant keywords to a user during web crawling, wherein the relevant keywords describe multiple attributes of a term that has an ambiguous meaning, and wherein the user is afforded an ability to specify keywords that have a minus degree of significance to a meaning intended by the user for web crawling; and narrowing down crawling objects by eliminating user-specified keywords that have a minus degree of significance, thereby eliminating ambiguity of a term being searched.” *Ishikawa* is cited for this teaching on col. 1, lines 53-67; col. 8, line 65 – col. 9, line 6; and col. 8, lines 13-16. However, the cited passages describe a method for ranking documents according to how often they use a keyword (see *Ishikawa*, col. 1, lines 22-23). Thereafter, documents can be ranked by both the parent page (the calling page) and the child page (the called page), as described at col. 8, line 65 – col. 9, line 6; and col. 8, lines 13-16 of *Ishikawa*. There is no teaching or suggestion of “eliminating ambiguity of a term being searched” by “narrowing down crawling objects by eliminating user-specified keywords that have a minus degree of significance,” as claimed and supported.

Additional comments

On pages 10-11, the Examiner states that in Applicant’s March 29, 2006 Amendment A, “the features upon which applicant relies (i.e., reading an HTML document of a web page as an analyzing object, conducting a temporary block analysis based on a description of HTML tags of

the HTML document) are not recited in the rejected claim(s).” Applicants disagree. The following is a copy of a relevant claim and arguments associated therewith, as found in the March 29, 2006 Amendment A:

Claim 8. An information search method for crawling a web site via a network using a computer, said method comprising the steps of:

acquiring a web page as initial information and storing source code into a storage device;
reading the source code of said web page from said storage device;
conducting a structure analysis of said web page, wherein the structure analysis includes the steps of:

reading an HTML document of a web page as an analyzing object,

conducting a temporary block analysis based on a description of HTML tags of the HTML document,

using the HTML tags to temporarily divide the HTML document into blocks, and
identifying unnecessary information elements in the HTML document, wherein the unnecessary information elements are plural information elements that include an OBJECT_IMAGE having a same Uniform Resource Locator (URL), wherein the OBJECT_IMAGE describes a type of media used to display the HTML document;
[[and]]

storing a result of the analysis into said storing device;

calculating a degree of significance of a web site linking from said web page, based on the result of said structure analysis stored in said storage device; and

accessing the web site depending on the calculated degree of significance to acquire contents thereof, and storing them into said storage device.

With regards to amended Claim 8, the cited art does not teach or suggest “conducting a structure analysis of said web page, wherein the structure analysis includes the steps of: reading an HTML document of a web page as an analyzing object, conducting a temporary block analysis based on a description of HTML tags of the HTML document (as supported in the present specification at paragraph [0075]), using the HTML tags to temporarily divide the HTML document into blocks, and identifying unnecessary information elements in the HTML

document, wherein the unnecessary information elements are plural information elements that include an OBJECT_IMAGE having a same Uniform Resource Locator (URL), (as supported at paragraph [0082]) wherein the OBJECT_IMAGE describes a type of media used to display the HTML document (as supported at paragraph [0059]).”

The argued and claimed language is identical, and thus Examiner’s response appears to be unfounded.

CONCLUSION

Note that the present amendment presents no new subject matter. As the cited art does not teach or suggest all of the presently claimed features, Applicants now respectfully request a Notice of Allowance for all pending claims.

No extension of time for this response is believed to be necessary. However, in the event an extension of time is required, that extension of time is hereby requested. Please charge any fee associated with an extension of time as well as any other fee necessary to further the prosecution of this application to **IBM CORPORATION DEPOSIT ACCOUNT No. 09-0461**.

Respectfully submitted,



James E. Boice
Registration No. 44,545
DILLON & YUDELL LLP
8911 North Capital of Texas Highway
Suite 2110
Austin, Texas 78759
512.343.6116

ATTORNEY FOR APPLICANT(S)



RECEIVED MAY 28 2004

JP920020109451
(W)

UNITED STATES PATENT AND TRADEMARK OFFICE

UNDER SECRETARY OF COMMERCE FOR INTELLECTUAL PROPERTY AND
DIRECTOR OF THE UNITED STATES PATENT AND TRADEMARK OFFICE

MAY 21, 2004

PTAS



102590634A

IBM CORPORATION
GREGORY M. DOUDNIKOFF
PO BOX 12195
RES. TRI. PARK, NC 27709

UNITED STATES PATENT AND TRADEMARK OFFICE
NOTICE OF RECORDATION OF ASSIGNMENT DOCUMENT

THE ENCLOSED DOCUMENT HAS BEEN RECORDED BY THE ASSIGNMENT DIVISION OF THE U.S. PATENT AND TRADEMARK OFFICE. A COMPLETE MICROFILM COPY IS AVAILABLE AT THE ASSIGNMENT SEARCH ROOM ON THE REEL AND FRAME NUMBER REFERENCED BELOW.

PLEASE REVIEW ALL INFORMATION CONTAINED ON THIS NOTICE. THE INFORMATION CONTAINED ON THIS RECORDATION NOTICE REFLECTS THE DATA PRESENT IN THE PATENT AND TRADEMARK ASSIGNMENT SYSTEM. IF YOU SHOULD FIND ANY ERRORS OR HAVE QUESTIONS CONCERNING THIS NOTICE, YOU MAY CONTACT THE EMPLOYEE WHOSE NAME APPEARS ON THIS NOTICE AT 703-308-9723. PLEASE SEND REQUEST FOR CORRECTION TO: U.S. PATENT AND TRADEMARK OFFICE, ASSIGNMENT DIVISION, BOX ASSIGNMENTS, CG-4, 1213 JEFFERSON DAVIS HWY, SUITE 320, WASHINGTON, D.C. 20231.

RECORDATION DATE: 10/27/2003

REEL/FRAME: 014642/0595
NUMBER OF PAGES: 3

BRIEF: ASSIGNMENT OF ASSIGNOR'S INTEREST (SEE DOCUMENT FOR DETAILS).
DOCKET NUMBER: T81/503

ASSIGNOR:
NOMIYAMA, HIROSHI

DOC DATE: 08/04/2003

ASSIGNOR:
IWA0, TOSHITAKA

DOC DATE: 08/08/2003

ASSIGNEE:
INTERNATIONAL BUSINESS MACHINES
CORPORATION
NEW ORCHARD ROAD
ARMONK, NEW YORK 10504

SERIAL NUMBER: 10621474
PATENT NUMBER:

FILING DATE: 07/17/2003
ISSUE DATE:

TITLE: INFORMATION SEARCH SYSTEM, INFORMATION SEARCH METHOD, HTML DOCUMENT
STRUCTURE ANALYZING METHOD, AND PROGRAM PRODUCT



US 20030046276A1

(19) **United States**(12) **Patent Application Publication** (10) Pub. No.: **US 2003/0046276 A1**
(43) Pub. Date: **Mar. 6, 2003**
Gutierrez et al.(54) **SYSTEM AND METHOD FOR MODULAR
DATA SEARCH WITH DATABASE TEXT
EXTENDERS**

(52) U.S. Cl. 707/3

(75) Inventors: **Arnold M. Gutierrez**, Leander, TX
(US); **Kevin R. Holubar**, Austin, TX
(US); **Shannon James Kerlick**, Cedar
Park, TX (US); **Dan Jeffrey**
Mandelstein, Austin, TX (US)(57) **ABSTRACT**Correspondence Address:
Joseph T. Van Leeuwen
P.O. Box 81641
Austin, TX 78708-1641 (US)

A system and method for searching a database from a computer network is provided. A client computer sends a search request to a search engine. The search engine prepares a database request. The preparation may include converting the client's query into a structured query language command. The search engine sends the database request to one or more servers that include database management systems, such as IBM's DB2™. The servers receive the request and extract responsive data from the databases being managed by the database management system. The extracted data is returned to the search engine which is then formatted and returned to the client. In addition, the search engine may maintain a search index that includes a compilation of database indices that have been received from one or more servers. This search engine can be searched to gather results responsive to a client request.

(73) Assignee: **International Business Machines Corporation**, Armonk, NY

(21) Appl. No.: 09/947,872

(22) Filed: Sep. 6, 2001

Publication Classification(51) Int. Cl.⁷ G06F 7/00